

# APPENDIX A

## TECHNICAL NOTES

### General Information

The data in this report come from many sources, including surveys conducted by NSF, other Federal agencies and by non-Federal sources. The data reflect many methods of collection, such as universe surveys, sample surveys, and compilations of administrative records. Users should take great care when comparing data from different sources. Data often will not be strictly comparable due to differences in definitions, survey procedures, phrasing of questions, and so forth.

Survey accuracy is determined by the joint effects of “sampling” and “nonsampling” errors. Sampling errors arise because estimates based on a sample will differ from the figures that would have been obtained if a complete census had been taken.

All surveys, whether universe or sample, are also subject to nonsampling errors, which can arise from design, reporting, and processing errors as well as errors due to faulty response or nonresponse. These nonsampling errors include respondent-based events, such as some respondents interpreting questions differently from other respondents; respondents making estimates rather than giving actual data; and respondents being unable or unwilling to provide complete, correct information. Errors can also arise during the processing of responses, such as recording and keying errors.

### Racial/Ethnic Information

Data collection and reporting of the race/ethnicity of individuals pose several additional problems. First, both the naming of population subgroups and their definitions have often changed over time. Because this report draws on data from many sources, different terminology may have been used to obtain the various statistics presented here. Efforts have been made to maintain consistency throughout this text, but in some data reporting, it has been necessary to use distinct terminology that does not match other compilations.

Second, many of the groups of particular interest are quite small, so that it is difficult to measure them accurately without universe surveys. In some instances

sample surveys may not have been of sufficient scope to permit calculation of reliable racial/ethnic population estimates, so that results are not shown for all groups. In addition, the reader is cautioned that it is easy to overlook or minimize the heterogeneity within subgroups when only a single statistic is reported for the total racial/ethnic group.

### Information About Persons With Disabilities

The data on persons with disabilities in science and engineering are seriously limited for several reasons. First, the operational definitions of “disability” vary and include a wide range of physical and mental conditions. Different sets of data have used different definitions and thus are not totally comparable. The Americans with Disabilities Act of 1990 (ADA) encouraged progress toward standard definitions. Under the ADA, an individual is considered to have a disability if the person has a physical or mental impairment that substantially limits one or more of the major life activities, has a record of such impairment, or is regarded as having such an impairment. The ADA also contains definitions of specific disabilities. (See Chapter 2, p. 22.)

Second, data about disabilities frequently are not included in comprehensive institutional records (e.g., in registrars’ records in institutions of higher education). If included at all in institutional records, such information is likely to be kept only in confidential files at an office responsible for providing special services to students. Institutions are unlikely to have information regarding any persons with disabilities who have not requested special services. In the case of elementary/secondary school programs receiving funds to provide special education, however, counts for the entire student population identified as having special needs are centrally available.

The third limitation on information on persons with disabilities gathered from surveys is that it often is obtained from self-reported responses. Typically, respondents are asked if they have a disability and to specify

what kind of disability it is. Resulting data, therefore, reflect individual perceptions, not objective measures.

Finally, data on persons with disabilities are often derived from sample surveys whose main purpose is to derive estimates for a full population. Deriving estimates for any phenomenon that is applicable to a small proportion of the total is particularly difficult, especially when the sampling procedures do not have a way to “oversample” cases providing the characteristic of interest. Because persons with disabilities constitute a relatively small portion of the population and because oversampling is not possible, sample sizes may not be sufficiently large to permit calculation of reliable estimates.

An example in which these factors come together can be seen in the attempt to provide estimates of the proportion of the undergraduate student population with disabilities. Self-reported data from the undergraduate student population, queried on a survey to ascertain patterns of student financial aid, suggest that about 10 percent of the undergraduate population report having some disability; estimates from population surveys of higher education institutions, in contrast, place the estimate much lower, between 1 and 5 percent. Whether this discrepancy is the result of self-perception, incomplete reporting, nonevident disabilities, or differing definitions is difficult to ascertain.

Therefore, although considerable information is available on persons with disabilities and their status in the educational system and in the science and engineering workforce, it is often not possible to compare the numbers of persons with disabilities from different sources.

## Primary Non-NSF Sources

The following non-NSF sources were used for data tables in this report.

### **Survey of Income and Program Participation**

#### **Contact:**

Current Population Reports  
Bureau of the Census  
U.S. Department of Commerce  
Washington, DC 20233  
Tel: (301) 763-8300

The Survey of Income and Program Participation conducted by the Census Bureau provides information on the economic situation of households and persons in the United States. The survey collects data on basic social and demographic characteristics of persons in households, labor force activity, type and amount of income, participation status in various programs, and various supplementary modules, for ex-

ample, work history, health characteristics (including disability), assets and liabilities, and education and training.

A combined sample from the 1992 and 1993 panels of the Survey of Income and Program Participation provides the latest available data on the disability status of the noninstitutionalized population of the United States. A supplement containing an extensive set of questions about disability status was included as part of the ninth wave of the 1992 panel and the sixth wave of the 1993 panel. Both of these waves were fielded between September and December 1994. The total sample size for this study was approximately 40,000 interviewed households.

The disability supplements that have been asked in SIPP were designed to be consistent with the ADA definition of disability. The supplements obtain information on the ability to perform specific functional activities (seeing, hearing, having one's speech understood, lifting and carrying, climbing stairs, and walking); certain ADLs or activities of daily living (getting around inside the home, getting in and out of a bed or chair, bathing, dressing, eating, and toileting), and certain IADLs or instrumental activities of daily living (going outside the home, keeping track of money and bills, preparing meals, doing housework, and using the telephone). The survey also collects information on the use of such special aids as wheelchairs and canes, the presence of certain conditions related to mental functioning, and the ability to work at a job or business.

## **National Assessment of Educational Progress**

### **Contact:**

National Center for Education Statistics  
U.S. Department of Education  
555 New Jersey Avenue, NW  
Washington, DC 20208-5653  
Tel: (202) 219-1761

The National Assessment of Educational Progress (NAEP) is sponsored by the National Center for Education Statistics (NCES) and has been conducted since 1983 by the Educational Testing Service. The overall goal of the project is to determine the Nation's progress in education. Accordingly, NAEP encompasses a series of national sample surveys designed to assess students in 10 subject areas such as reading, mathematics, science, writing, and history. Begun in 1969, NAEP was conducted annually through 1980; since 1980 the project has been conducted biennially. NAEP has surveyed the educational accomplishments of 9-, 13-, and 17-year-old students (and, in recent years, those in grades 4, 8, and 12 as well).

Since 1986, NAEP has included both main and long-term trend assessments. Both assessments use a

complex multistage stratified sample of schools that are selected to ensure adequate representation of schools with high enrollment of blacks and Hispanics. Both assessments historically excluded students with limited English proficiency and students receiving special education services whom school officials judged unable to respond meaningfully to the assessment either because the students had limited English proficiency or because they had a severe mental or physical disability.

Beginning with NAEP assessments in 1996, attempts were made to have more of the students who were classified as having disabilities or limited English proficiency included in the assessment. Accommodations were implemented for students who would have been excluded in the past. Spanish-speaking students classified as Limited English Proficiency were given the option of using a bilingual test booklet in mathematics in a portion of the sample. In addition, English-Spanish glossaries were provided at all three grades for designated science books. Other accommodations (such as earphones for the hearing impaired, signers for the deaf, magnifying equipment, and translators) were allowed if provided by the school and specified in the student's Individual Education Plan. A study to determine the impact of the revised inclusion rules and accommodations is taking place with the 1996 assessment. Students classified as disabled or with limited English proficiency who cannot be accommodated, either by the administrators of NAEP or by the schools, are excluded from the assessment.

The main assessments estimate student achievement at a cross-sectional point in time. The cross-sectional samples use innovations in assessment methodology and population definition. In 1996, data were collected from approximately 35,000 students in grades 4, 8, and 12 for the science assessments and from approximately 33,000 students in grades 4, 8, and 12 for the mathematics assessments. Data were also collected from these students' principals and a sample of their teachers.

The long-term trend assessments estimate the current status of achievement using the same sampling and assessment methodology used in previous years. In 1996, approximately 15,000 students ages 9, 13, and 17 were tested in mathematics and in science.

Performance data are reported for the Nation, and for various subgroups categorized by such variables as region, gender, race/ethnicity, parental education, type of school, and type and size of community. Also initiated in 1983 was the reporting of performance data by scaled proficiency levels. Beginning with the 1990 assessment, National Assessment Governing Board established three reporting levels for reporting NAEP results: basic, pro-

ficient, and advanced. Currently, NAEP is conducted every other year in even-numbered years.

### **American College Testing Program**

#### **Contact:**

The American College Testing Program  
2201 North Dodge Street  
P.O. Box 168  
Iowa City, IA 52243  
Tel: (319) 337-1510

The American College Testing (ACT) Assessment is taken by college-bound high school students who request that the results be sent to designated colleges and scholarship boards. The ACT is designed to measure educational development in the areas of English, mathematics, social studies, and natural sciences. The test results are used in part to help predict how well students might perform in college. In 1994, approximately 892,000 students took the ACT examinations.

ACT standard scores are reported for each subject area on a scale from 1 to 36. A composite score is obtained by taking the simple average of the four standard scores and is an indication of a student's overall academic development across the four subject areas.

Since the 1984-1985 school year, national norms have been based on the most recent ACT test scores available from all students taking the test and who are scheduled to graduate in the spring of the year.

It should be noted that college-bound students who take the ACT Assessment are not, in some respects, representative of college-bound students nationally. First, students who live in the Midwest, South, Rocky Mountains, and Plains regions are overrepresented among ACT-tested students compared with college-bound students nationally. Second, ACT-tested students tend to enroll in public colleges and universities more frequently than do college-bound students nationally.

### **Scholastic Assessment Test (SAT)**

#### **Contact:**

College Entrance Examination Board  
Educational Testing Service  
Princeton, NJ 08541  
Tel: (609) 771-7600

The Admissions Testing Program of the College Board comprises a number of college admissions tests, including the Scholastic Assessment Test (SAT). The SAT is taken by students who need the results to apply to a particular college or university or scholarship board. High school students participate in the testing program as sophomores, juniors, or seniors—some more than once during these 3 years. If they have taken the tests more than once, only the most recent scores are tabulated.

The SAT reports subscores in the areas of mathematics and verbal ability. Students may also elect to take Achievement Tests in any of 21 subject areas; these exams are generally taken by students who are applying to the more competitive schools. In 1994, approximately 1.1 million students took the SAT examination, and more than 200,000 took at least one Achievement Test.

In 1987 the College Board initiated a review of the Admissions Testing Program and made significant changes in the SAT Program in 1993–1994. Through the January 1994 test administration, SAT Program tests included the SAT, the Test of Standard Written English (TSWE), and the Achievement Tests. Beginning in March 1994, the SAT Program was revised into two formats: the SAT I: Reasoning Test (the mathematical and verbal sections, with revisions beginning in March 1994) and SAT II: Subject Tests (formerly known as the Achievement Tests, with the revisions beginning in May 1994).

The SAT results are not representative of high school students or college-bound students nationally since the sample is self-selected. In addition, public colleges in a number of states require that students applying for admission submit ACT scores rather than SAT scores; thus, the proportion of students taking the SAT in some states is very low.

### **The Integrated Postsecondary Education Data System Survey: Fall Enrollment, Completions and Institutional Characteristics**

#### **Contact:**

National Center for Education Statistics  
U.S. Department of Education  
555 New Jersey Avenue, NW  
Washington, DC 20208-5652  
Tel: (202) 219-1373

The Integrated Postsecondary Education Data System (IPEDS) began in 1986 as a supplement to and replacement for the Higher Education General Information Survey (HEGIS), which began in 1966. HEGIS was an annual survey of institutions listed in the current NCES *Education Directory of Colleges and Universities*; IPEDS surveys all postsecondary institutions, including universities and colleges and the institutions that offer technical and vocational education. The higher education portion is a census of accredited 2- and 4-year colleges, whereas technical and vocational schools are surveyed on a sample basis.

IPEDS consists of several integrated components that obtain information on types of institutions where postsecondary education is available, student participants, programs offered and completed, and the human and financial resources involved in the delivery of postsecondary education. The components of

IPEDS include surveys of institutional characteristics; fall enrollment of students, including their age and residence; fall enrollment in occupationally specific programs; completions; finance; staff; salaries of full-time instructional faculty; and academic libraries.

The IPEDS Institutional Characteristics Survey provides the basis for the universe of institutions reported in the *Education Directory of Colleges and Universities*. The universe includes institutions that met certain accreditation criteria and offered at least a 1-year program of college-level studies leading toward a degree. Each fall, institutions listed in the previous year's directory are asked to update information on the characteristics of their schools.

The IPEDS Completions Survey replaces and extends the HEGIS Degrees and Other Formal Awards Conferred Survey. The Completions Survey is administered to a census of institutions offering degrees at the bachelor's level and above, all 2-year institutions, and a sample of less-than-2-year institutions.

The IPEDS Fall Enrollment Survey replaces and extends the previous HEGIS surveys of institutions of higher education.

### **The National Postsecondary Student Aid Survey**

#### **Contact:**

National Center for Education Statistics  
U.S. Department of Education  
555 New Jersey Avenue, NW  
Washington, DC 20208-5652  
Tel: (202) 219-1839

The National Postsecondary Student Aid Study (NPSAS) was established by NCES to collect information concerning financial aid allocated to students enrolled in U.S. postsecondary institutions. NPSAS was first administered in the fall of the 1986–1987 academic year. NCES conducted subsequent cycles of NPSAS for the 1989–1990, 1992–1993, and 1995–1996 school years. The 1989–1990 cycle contained enhancements to the methodology used in the 1987 cycle. Estimates from the 1996 NPSAS sample are generally comparable to those from the 1993 and 1990 samples but not to those from the 1987 sample.

The 1995–1996 survey gathered information from about 60,000 undergraduate and graduate students selected from registrar lists of enrollees at about 800 postsecondary institutions. The sample included students who did and did not receive financial aid, as well as students' parents. Student information, such as field of study, educational level, and attendance status (part-time or full-time), was obtained from registrar records. Types and amounts of financial aid and family financial characteristics were abstracted from school

financial aid records. Parents of students were also sampled to compile data concerning family composition and parental financial characteristics. Biennial follow-up data collections are expected.

## Primary NSF Sources

The following NSF sources were used for data tables in this publication. Published data tables from these surveys may be accessed on the NSF Web page (<http://www.nsf.gov/sbe/srs>). In addition, researchers may access the data directly through the SESTAT or WebCASPAR database systems on the Web.

## Survey of Earned Doctorates

The Survey of Earned Doctorates (SED) has been conducted annually since 1957. Until 1996, it was conducted under contract with the National Research Council of the National Academy of Sciences, for the National Science Foundation, the U.S. Department of Education, the National Endowment for the Humanities, the National Institutes of Health, and the U.S. Department of Agriculture. This is a census survey of all recipients of research doctoral degrees such as PhD or D.Sc.; it excludes the recipients of first-professional degrees such as J.D. or M.D. Therefore, SED data are restricted to research doctorates.

Data for the SED are collected directly from individual doctorate recipients contacted through graduate deans at all U.S. universities. The recipients are asked to provide information on the field and specialty of their degree, as well as their personal educational history, selected demographic data, and information on their postgraduate work and study plans. Approximately 95 percent of the annual cohort of doctorate recipients respond to the questionnaire, which is distributed through the cooperation of the graduate deans at institutions awarding doctorates.

Partial data from public sources, such as field of study, are added to the file for nonrespondents. No imputations are made, however, for nonresponse for data not available elsewhere, such as race/ethnicity information. The data for a given year include all doctorates awarded in the 12-month period ending on June 30 of that year.

## Survey of Graduate Students and Postdoctorates in Science and Engineering

The data collected in the fall 1995 Survey of Graduate Students and Postdoctorates in Science and Engineering (GSS) represent national estimates of graduate enrollment and postdoctoral employment at the beginning of academic year 1995–1996 in all academic institutions in the United States that offer doctorate or

master's degree programs in any science or engineering field. Included are data for all branch campuses, affiliated research centers, and separately organized components, such as medical or dental schools, schools of nursing, public health. The survey universe consisted of 722 reporting units at 602 graduate institutions. Data are collected at the academic department level.

Available information includes full-time graduate students by source and mechanism of support, including data on women and first-year students enrolled full time; part-time graduate students by sex; and citizenship and racial/ethnic background of all graduate students. In addition, detailed data on postdoctorates are available by source of support, sex, and citizenship, including separate data on those holding first-professional doctorates in the health fields; summary information on other doctorate nonfaculty research personnel is also included.

The National Science Foundation has collected data on graduate science and engineering enrollment and postdoctoral appointees since 1966. From fall 1966 through fall 1971, data from a limited number of doctorate-granting institutions were collected through the NSF Graduate Traineeship Program, which requested data only on those science and engineering fields supported by NSF. Beginning with the fall 1972 survey, this data-collection effort was assigned to the Universities and Nonprofit Institutions Studies Group of SRS. It was gradually expanded during the period 1972–1975 to include additional science and engineering fields as well as all institutions known to have programs leading to the master's or doctorate degree. Because of this expansion, data for 1974 and earlier years are not strictly comparable with 1975 and later data.

## NSF's SESTAT Data System

In the 1990s, NSF redesigned its data system about scientists and engineers. Termed SESTAT, the new data system integrates data from three NSF surveys (the Survey of Doctorate Recipients, the National Survey of College Graduates, and the National Survey of Recent College Graduates). The integration of the SESTAT surveys requires complementary sample populations and reference periods, matching survey questions and procedures, as well as weighting adjustments for any overlapping populations.

The surveys provide data on educational background, occupation, employment, and demographic characteristics. These surveys are of individuals and have a combined sample size of about 105,000, covering a population of about 12 million scientists and engineers. SESTAT defines scientists and engineers as those who either received a college degree (bachelor's level or higher) in a science or engineering field or

who work as a scientist or engineer. Each of the three surveys that make up the SESTAT data system collects new data every 2 years. The data reported in this publication were collected in 1995.

SESTAT has as its target population residents of the United States with a baccalaureate degree or higher who, as of the study's reference period, were noninstitutionalized, age 75 or less, and either trained as or working as a scientist or engineer (S&E). A baccalaureate-or-higher degree is a bachelor's, master's, doctorate, or professional degree. To meet the S&E requirement, the U.S. resident had to (1) have at least one baccalaureate-or-higher degree in an S&E field or (2) have a baccalaureate-or-higher degree in a non-S&E field and work in an S&E occupation as of the reference week. For the 1995 SESTAT, the reference period was the week of April 15, 1995.

Some elements of SESTAT's desired target population were not included within the target populations of any of the three SESTAT component surveys. Bachelor's- and master's-level S&E-trained personnel missing from the survey frames are predominately

- residents whose S&E bachelor's and/or master's degrees were received prior to April 1990 or from a foreign institution, who resided outside the United States on April 1, 1990, but not as U.S. armed forces stationed abroad;
- residents with no baccalaureate or higher degree in any field as of April 1, 1990, who were awarded an S&E degree after June 1994 by a U.S. institution or after April 1990, by a foreign institution.

Doctorate level S&E-trained personnel missing from the survey frames are predominately

- residents with S&E doctorates received after June 1994 or from a foreign institution, with no baccalaureate-or-higher degree in any field as of April 1, 1990, and no bachelor's or master's S&E degree received from a U.S. institution between April 1, 1990, and June 1994;
- residents with S&E doctorates received after June 1994 or from a foreign institution but with no bachelor's or master's S&E degree received from a U.S. institution between April 1, 1990, and June 1994, who resided outside the United States on April 1, 1990, but not as U.S. armed forces stationed abroad.

SESTAT classifies the following broad occupation categories as S&E occupations: computer and mathematical scientists, life and related scientists, physical and related scientists, social and related scientists, and engineers. Postsecondary teachers are included within each of these groups. The following are considered non-S&E occupations: top and midlevel managers; teachers, except S&E postsecondary teachers; technologists, including computer programmers; people in health and related occupations, social services and related occupations, sales and marketing occupations, and other non-S&E occupations (for example, artists, broadcasters, editors, entertainers, public relations specialists, writers, clerical and administrative support personnel, farmers, foresters, fishermen, lawyers, judges, librarians, archivists, curators, actuaries, food service, historians (except science and technology), construction trades people, mechanics and repairers, and those involved in precision/production occupations, operators (for example, machine set-up, machine operators and tenders, fabricators, assemblers) and related occupations, transportation/material moving occupations and protective and other service occupations.)